

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平11-249867

(43)公開日 平成11年(1999) 9月17日

(51)Int.Cl.<sup>6</sup>  
G 0 6 F 3/16識別記号  
3 4 0F I  
G 0 6 F 3/16

3 4 0 X

3 4 0 Z

13/00

3 5 1

13/00

3 5 1 B

3 5 4

3 5 4 D

審査請求 未請求 請求項の数6 O L (全 11 頁)

(21)出願番号 特願平10-48180

(22)出願日 平成10年(1998) 2月27日

(71)出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(71)出願人 591230295

エヌティティエレクトロニクス株式会社

東京都渋谷区桜丘町20番1号

(72)発明者 並木 育夫

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

(72)発明者 林 弘道

東京都新宿区西新宿三丁目19番2号 日本

電信電話株式会社内

(74)代理人 弁理士 伊東 忠彦

最終頁に続く

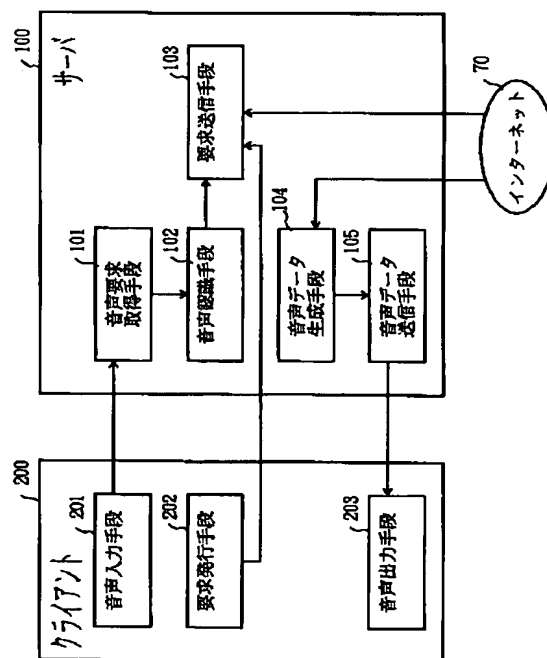
(54)【発明の名称】 音声ブラウザシステム

(57)【要約】

【課題】 視覚障害者であってもWWW情報を取得することが可能な音声ブラウザシステムを提供する。

【解決手段】 本発明は、クライアントから音声入力による要求を取得する音声要求取得手段と、音声要求取得手段より入力された音声による要求を音声認識する音声認識手段と、音声認識手段により認識された結果に基づいてクライアントから指定されたURLに対して、要求をインターネットに送信する要求送信手段と、インターネットから取得した応答から音読テキストを抽出して音声データに変換し、音声合成する音声データ生成手段と、音声データ生成手段により生成された音声データをクライアントに送信する音声データ送信手段とを有するサーバと、ユーザの音声による要求を入力する音声入力手段と、サーバから取得した結果よりURLを抽出し、該URLに基づいて該サーバにHTMLファイルを要求する要求発行手段と、サーバから送信された音声データを出力する音声出力手段とを有するクライアントとを有する。

本発明の原理構成図



### 【特許請求の範囲】

【請求項1】 コンピュータとインターネットワークからなるクライアント／サーバシステムにおいて、サーバに蓄積されている情報をクライアント側で音声により出力する音声ブラウザシステムであって、

前記サーバは、

クライアントから音声入力による要求を取得する音声要求取得手段と、

前記音声要求取得手段より入力された音声による要求を音声認識する音声認識手段と、

前記音声認識手段により認識された結果に基づいて前記クライアントから指定されたURLに対して、前記要求を前記インターネットに送信する要求送信手段と、

前記インターネットから取得した応答から音読テキストを抽出して音声データに変換し、音声合成する音声データ生成手段と、

音声データ生成手段により生成された音声データを前記クライアントに送信する音声データ送信手段とを有し、前記クライアントは、

ユーザの音声による要求を入力する音声入力手段と、

前記サーバから取得した結果よりURLを抽出し、該URLに基づいて該サーバにHTMLファイルを要求する要求発行手段と、

前記サーバから送信された音声データを出力する音声出力手段とを有することを特徴とする音声ブラウザシステム。

【請求項2】 前記音声認識手段は、

前記クライアントから入力された音声データの要求に類似する候補からなるリンク項目リストと、

前記リンク項目リストを形態素解析する形態素解析手段を含む請求項1記載の音声ブラウザシステム。

【請求項3】 前記音声データ生成手段は、

前記インターネットから取得した応答を前記クライアントが希望する言語に翻訳する翻訳手段を含む請求項1記載の音声ブラウザシステム。

【請求項4】 前記音声出力手段は、

前記サーバから受信した音声データが、音楽コンテンツまたは、朗読コンテンツである場合に、再生を行う再生手段を含む請求項1記載の音声ブラウザシステム。

【請求項5】 前記再生手段は、

一時停止モード、再生モード、停止モードを有する請求項4記載の音声ブラウザシステム。

【請求項6】 前記再生手段は、

前記音声データの再生中に、音声情報による割り込み出力を行う割り込み手段を含む請求項4乃至5記載の音声ブラウザシステム。

### 【発明の詳細な説明】

#### 【0001】

【発明の属する技術分野】本発明は、音声ブラウザシステムに係り、特に、コンピュータとネットワークからな

るクライアント／サーバ構成のシステム、とりわけインターネットにおけるWorld Wide Web（以下、単にWWWと記す）システムのサーバに対し、クライアント端末のマイクから音声で入力し、サーバに蓄積している情報を音声で出力する音声ブラウザシステムに関する。

#### 【0002】

【従来の技術】周知のように、WWWシステムとして、サーバ及びクライアントのハードウェア、ソフトウェアがネットワーク上に適切に構成されている場合に、クライアント端末上にインストールしたNetscape Navigatorなどのブラウザを使用することによって、サーバに格納したテキストやイメージの情報をクライアント画面上に表示して閲覧することが可能である。

【0003】このシステムでは、画面上の特定の情報をマウスなどで選択すると、これと関連付けられた（以下では、これをリンクが張られたと言い、先の特定の情報をリンク項目と言う）情報にアクセスし、画面上に表示し閲覧することが可能である。これらのサービスは、情報を視覚によって享受することが前提であり、画面に目を向けないと享受できない、あるいは、視覚障害者は全く享受できないという欠点がある。これを解決する方法として、最近の音声認識技術と音声合成技術を使用することによって、マイクから音声で入力し、音声合成で出力することが可能である。例えば、音声で「首相官邸」と入力すれば、「首相官邸」の情報にアクセスし、クライアント端末のスピーカから、テキスト部分を合成音で出力することができる。

#### 【0004】

【発明が解決しようとする課題】しかしながら、上記従来の方法では、周知のWWW情報には、長い文章や、リンク箇所が随所に10個も20個もある場合がある、カラーのイメージ情報がテキストに混在することは勿論のこと、動画とリンクが張られたり、視覚に訴える情報がふんだんに使用されるのが実情である。このような情報を視覚障害者にいかに出力するかという問題が存在する。

【0005】本発明は、上記の点に鑑みなされたもので、視覚障害者であってもWWW情報を取得することが可能な音声ブラウザシステムを提供することを目的とする。

#### 【0006】

【課題を解決するための手段】図1は、本発明の原理構成図である。本発明（請求項1）は、コンピュータとインターネットワークからなるクライアント／サーバシステムにおいて、サーバ100に蓄積されている情報をクライアント200側で音声により出力する音声ブラウザシステムであって、サーバ100は、クライアント200から音声入力による要求を取得する音声要求取得手段201と、音声要求取得手段201より入力された音声による要求を音声認識する音声認識手段102と、音声

認識手段102により認識された結果に基づいてクライアント200から指定されたURLに対して、要求をインターネット70に送信する要求送信手段103と、インターネット70から取得した応答から音読テキストを抽出して音声データを変換し、音声合成する音声データ生成手段104と、音声データ生成手段104により生成された音声データをクライアント200に送信する音声データ送信手段105とを有し、クライアント200は、ユーザの音声による要求を入力する音声入力手段201と、サーバ100から取得した結果よりURLを抽出し、該URLに基づいて該サーバに、インターネット70上に公開されているHTMLファイルを要求する要求発行手段202と、サーバ100から送信された音声データを出力する音声出力手段203とを有する。

【0007】本発明（請求項2）は、音声認識手段102において、クライアント200から入力された音声データの要求に類似する候補からなるリンク項目リストと、リンク項目リストを形態素解析する形態素解析手段を含む。本発明（請求項3）は、音声データ生成手段104において、インターネット70から取得した応答をクライアントが希望する言語に翻訳する翻訳手段を含む。

【0008】本発明（請求項4）は、音声出力手段203において、サーバ100から受信した音声データが、音楽コンテンツまたは、朗読コンテンツである場合に、再生を行う再生手段を含む。本発明（請求項5）は、再生手段において、一時停止モード、再生モード、停止モードを有する。

【0009】本発明（請求項6）は、再生手段において、音声データの再生中に、音声情報による割り込み出力を行う割り込み手段を含む。上記のように、本発明では、インターネット上に公開されているHTML(Hyper Text Markup Language)形式のファイルを市販のWebブラウザを通して視覚情報から音声情報に変換し、ユーザに提供することを可能にしたシステムである。また、クライアント側で情報を取得する場合において、音声を用いることで視覚障害者の操作も可能とする。

#### 【0010】

【発明の実施の形態】図2は、本発明が適用されるシステムの構成を示す。同図に示すシステムは、高速なネットワーク上に各処理用エンジンを配置し、負荷分散を行うことにより高速なレスポンスをクライアント端末10で実現可能とするものである。同図におけるシステムでは、大別して2つのシステムに分けられる。

【0011】まず、第1に同図におけるワークステーション20、30、40で構成されるフロント・エンドとしての処理部である。ワークステーション20、30は、翻訳サービスを提供するためのシステムである。ワークステーション40は、一般的にインターネットでよく利用される機能で、本システムでは主に、データのキ

ャッシングや漢字コードの変換で利用する。

【0012】第2に、バック・エンドとして利用されるワークステーション50、60である。ワークステーション50は、ユーザからクライアント端末10に入力された音声データを、候補リスト（リンク項目のリスト）と比較照合を行い、適切な項目を選出する。また、ユーザが入力した音声データとリンク項目が全く同一である必要性を解消するためにリンク項目に対して形態素解析処理を施す。そうすることにより、ユーザからはリンク項目の断片を入力するだけでも適切なリンク項目を推量し、選出する。ワークステーション60は、クライアント端末10で抽出されたテキスト情報を、（言語種別のパラメータ付きで）受信し、音声データを生成する音声合成エンジンと呼ぶものである。

【0013】次に、クライアント端末10から送信されたHTMLファイル要求は、ワークステーション20上で動作中にプロキシ・サーバを経由し、ワークステーション40のプロキシ・サーバを使って外部インターネット70に転送される。インターネット70から返送されてきた応答データ（HTMLファイル）は、ワークステーション40のプロキシ・サーバを経由して、ワークステーション20のプロキシ・サーバで各翻訳エンジンへデータの処理を依頼して（依頼のみ、処理結果はユーザの要求で随時クライアント端末10へ送信される）クライアント端末10へ転送する。クライアント端末10に届いたHTMLファイルは、解析され、表示されるべきテキスト情報をワークステーション60の音声合成エンジンへ送信し、音声データへ変換され、クライアント端末10で出力される。また、音声認識のために必要なリンク項目のリストは、ワークステーション50の形態素解析エンジンで形態素解析処理を施され、音声認識エンジンへ渡され、クライアント端末10から音声データの送信を待つ。

【0014】次にクライアント端末10で入力（指示）された音声データを音声認識エンジン50に送信し、音声認識エンジン50は、以前入力されていたリンク項目（テキスト情報）と比較照合され、適切な結果を得る。得られた結果をクライアント端末10に送信する。これにより、クライアント端末10は、リンク項目からURLを抽出し、インターネット70へプロキシサーバ20、40を介して、次の情報を取得する。

#### 【0015】

【実施例】以下に、本発明の実施例を図面と共に説明する。図3は、本発明の一実施例のシステム構成を示す。同図に示すシステムにおいて、図2と同様のものについては同一符号を付す。同図に示すシステムは、クライアント端末10、ワークステーション20、30、40、50から構成される。

【0016】クライアント端末10は、Webブラウザ11、音声入力用ボタン監視プログラム12、音声ブラ

ウザククライアント 13、音声入力用ボタン監視プログラム 12 に接続される音声入力開始用ボタン 15、音声ブラウザクライアント 13 に接続されるスピーカ 16 及びマイク 17 から構成される。ワークステーション 20 は、プロキシサーバ 21 の機能と、テキストを英語から日本語に翻訳する英日翻訳エンジン 22 を有する。

【0017】ワークステーション 30 は、テキストを日本語から英語に翻訳する日英翻訳エンジン 31 を有する。ワークステーション 40 は、プロキシサーバ 41 の機能を有する。ワークステーション 50 は、音声認識 I/F プログラム 51、音声認識エンジン 52、形態素解析エンジン 53 を有する。

【0018】ワークステーション 60 は、音声合成エンジン 61 を有する。Web ブラウザ 11 は、一般によく利用されるもので、例えば、Netscape Navigator 等がある。本実施例では、当該ブラウザを用いて説明する。当該 Web ブラウザ 11 は、インターネット 70 への窓口として必要情報を取得し、音声ブラウザクライアント 13 に渡す。また、Web ブラウザ 上への表示も行う。音声ブラウザクライアント 13 では、Web ブラウザ 11 から得られた情報を解析し、音読すべきテキスト情報をワークステーション 60 の音声合成エンジン 61 へ、音声認識すべきリンク項目の情報をワークステーション 50 の音声認識 I/F プログラム 51 へそれぞれ送信する。

【0019】クライアント端末 10 は、音声合成エンジン 61 から受信した音声データをローカル・ディスクに記録しながら、スピーカ 16 により再生する。また、ユーザからの入力は、音声入力開始用ボタン 15 を押下することによって、音声入力用ボタン監視プログラム 12 から音声ブラウザクライアント 13 に通知される。その通知を受けた音声ブラウザクライアント 13 は、マイク 17 から音声の録音を開始する。音声ブラウザクライアント 13 は、ユーザの音声入力開始用ボタン 15 の解放によって、録音の停止を行い、音声認識 I/F プログラム 51 に録音した音声データをワークステーション 60 の音声合成エンジン 61 に送信する。

【0020】クライアント端末 10 に接続れるワークステーション 20 のプロキシサーバ 21 は、インターネット 70 からの転送すべき情報 (HTML ファイル) を各翻訳エンジン (英日翻訳エンジン 22、日英翻訳エンジン 31) へ送信し、翻訳処理を行わせる。翻訳結果については、各翻訳エンジン 22、31 の記憶装置に記録しておき、ユーザから翻訳要求があった際にその翻訳結果をクライアント端末 10 に送信する。

【0021】ワークステーション 40 のプロキシサーバ 41 では、インターネット 70 からの情報を一部変換 (漢字コード等) したり、情報を一時的にキャッシュするなどの機能を有する。ワークステーション 50 の音声認識 I/F プログラム 51 は、音声ブラウザクライアント 1

3 から送信されるリンク項目を形態素解析エンジン 53 に入力し、出力された結果である分解された文字列の品詞を元に適切な再構成処理を行う。その結果を音声認識エンジン 52 へ登録し、音声ブラウザクライアント 13 から送信される録音された音声データを音声認識エンジン 52 に渡し、当該音声認識エンジン 52 において照合された結果を音声ブラウザクライアント 13 に返送する。

【0022】ワークステーション 60 の音声合成エンジン 61 は、音声ブラウザクライアント 13 で抽出された音読すべきテキスト情報を受信し、音声データを生成し、音声ブラウザクライアント 13 に返送する。図 4 は、本発明の一実施例の動作のシーケンスチャートである。まず、ユーザが音声入力開始用ボタン 15 を押下すると (ステップ 101)、音声入力用ボタン監視プログラム 12 から音声ブラウザクライアント 13 へその通知が転送される。音声ブラウザクライアント 13 は、マイク 17 から音声録音を開始し、終了を待つ。ユーザによりボタンが解放になると (ステップ 102)、音声入力用ボタン監視プログラム 12 から再度終了の通知が音声ブラウザクライアント 13 に報告される。その報告を受けた音声ブラウザクライアント 13 は、録音を停止し、ワークステーション 50 の音声認識 I/F プログラム 51 へその録音された音声データを送信する (ステップ 103)。

【0023】ワークステーション 50 の音声認識 I/F プログラム 51 は、当該データを受信し、音声認識エンジン 52 にそのデータを転送し (ステップ 104)、認識処理を依頼する。当該音声認識エンジン 52 により認識処理された結果を音声人気 I/F プログラム 51 が取得すると (ステップ 105)、当該認識結果が音声ブラウザクライアント 13 に転送される (ステップ 106)。音声ブラウザクライアント 13 は、結果のリンク項目から URL を得て、Web ブラウザ 11 に要求送信指示を発行する (ステップ 107)。

【0024】Web ブラウザ 11 は、指定された URL へデータ要求伝文をプロキシサーバ 21、41 を経由してインターネット 70 に送信する (ステップ 108)。プロキシサーバ 21 がプロキシサーバ 41 を経由して、当該要求に対応する応答をインターネット 70 から取得すると (ステップ 109)、当該応答を英日翻訳エンジン 22 または、日英翻訳エンジン 31 に送信する (ステップ 110)。

【0025】英日翻訳エンジン 22、日英翻訳エンジン 31 のいずれかがプロキシサーバ 21 により指示された処理を行い、その結果をプロキシサーバ 21 に返却する。これによりプロキシサーバ 21 は、応答データをクライアント端末 10 の Web ブラウザ 11 に送信する (ステップ 111)。クライアント端末 10 の Web ブラウザ 11 が応答データを受け取ると、音声ブラウザク

クライアント13に渡される(ステップ112)。このとき、ページがマルチフレームの場合は、Webブラウザ11から上記の処理を構成ビュー分繰り返す。音声ブラウザクライアント13では、取得した応答データを解析し、リンク項目と表示されるテキスト情報等を取得し、それぞれをワークステーション60の音声合成エンジン61及び、ワークステーション50の音声認識I/Fプログラム51に送信する(ステップ113)。その際、HTML解析を行い、イメージ(絵)情報に付加されている説明文等のテキスト情報がある場合は、その情報も

適切に処理し、音声合成エンジン61へ送信され、ユーザにイメージの子細をスピーカ16より音読で伝える。また、音声合成エンジン61に送信する場合は、言語別に1文単位で送信を行うことによって適切なサービス

(1文単位の巻き戻し/早送り等)をユーザに提供することを可能にしている。音声認識I/Fプログラム51には、リンク項目など動的に変化する認識候補リスト情報などを送信する(ステップ114)。なお、固定コマンドについては、毎回送信は行わないものとする。

【0026】音声認識I/Fプログラム51では、受信したリンク項目について1度形態素解析エンジン53により形態素解析処理を施し(ステップ115、116)、音声認識エンジン52へ固定コマンドと共に登録され(ステップ117)、ユーザからの音声データを待機する。なお、当該登録は、ユーザから送信される音声データとの比較照合に利用するために必要となる。

【0027】音声合成エンジン60は、ユーザからの音声合成し、合成された音声データを音声ブラウザクライアント13に送信する(ステップ118)。次に、具体的な例を用いて説明する。図5は、本発明の一実施例のユーザインターフェースを示す図であり、図6は、本発明の一実施例の音声ブラウザにより音読される例を示す図である。

【0028】同図は、Webブラウザ11のユーザインタフェースである。ページには、タイトル110、リンク項目130、本文120がテキスト情報で表示されている。これらの情報は、HTMLに従った文法で書かれたファイルから得られ、当該ファイルを音声ブラウザクライアント13側で解析を行い、タイトル、リンク項目、本文といった項目に分類する。そして、出力される情報は、図5に示すように音読される。適切なガイダンスを付加することによって、ユーザに詳細に情報を提供する。

【0029】また、入力時において、リンク項目を指示する場合には、リンク項目全文を音声ブラウザクライアント13に入力することなく、印象に残った単語だけの入力でアクセスが可能である。その仕組みについては、まず、形態素解析エンジン53にリンク項目全文を入力し、品詞分解された結果から再度組み合わせを行う。最小の要素である単語から複合語などを再構築していくこ

とによって、ユーザからの(単語から複合語までの)入力に対処できる。

【0030】図7は、本発明の一実施例のWebブラウザのマルチフレーム構成によって表示されているホームページの型であり、図8は、本発明の一実施例の音声ブラウザにより音読される例である。この場合には、音声ブラウザクライアント13では、複数のビューを持つことなどをHTMLファイルを解析することによって取得し、ユーザに音声で伝える。また、音読はビュー単位で行われるものとする。

【0031】図9は、本発明の一実施例のWebブラウザにイメージが掲載された例であり、図10は、本発明の一実施例の音声ブラウザにより音読される例である。図9に示すイメージ図は、視覚情報であるため、当該情報を音声で伝えるのは困難である。しかし、HTMLのタグ情報を解析することによって、イメージに説明を付加することによって音声ブラウザクライアント13側でイメージに対応する説明を抽出し、音声で読み上げることが可能とする。この場合は、HTML作成者がそのイメージ図に関する説明文をテキスト情報で追記することを条件とする。

【0032】次に、再生制御機能について説明する。この機能は、音読の読み上げ速度、音量、話者の性別などを音声の指示によってリアルタイムに変更できる機能である。読み上げ速度、話者の性別に関しては、音声データの再作成を音声合成エンジン61のパラメータを変更することによって行い、また、再生ポイントから優先的に再作成を行うことによってリアルタイムに変更を可能にした。音量については、システムのパラメータを変更することによって対処するものとする。

【0033】再制御機能の付加機能として、特定のサービス(現在時刻の通知等)の割り込み等も付加することが可能である。詳しくは、音読中に現在時刻を質問すると、システムに時刻を問い合わせ、一度音声合成エンジン61へ音声データの作成を要求し、作成完了と共に音読を一時中断し、現在時刻をアナウンスする。また、その後、中断中の音読を再開することによって実現できる。

【0034】また、HTMLファイル中に埋め込まれている音声データの再生制御も可能である。図11は、本発明の一実施例の音楽/朗読コンテンツ再生を行う場合の構成図である。この音声データは、音楽や朗読などの内容のもので再生する際には、再生可能な再生用ソフトウェア14をWebブラウザ11が自動的に起動する。その再生用ソフトウェア14を音声ブラウザクライアント13が制御することによって、一時停止や再生などの当該ソフトウェアが持つ従来の機能を音声で制御することが可能となる。

【0035】図12は、本発明の一実施例の音声データの再生制御のシーケンスチャートである。まず、音声ブ

10

20

30

40

50

ブラウザクライアント13からWebブラウザ11にURLの指示を発行すると(ステップ201)、Webブラウザ11は、当該指示をインターネット70に送信する。これによりWebブラウザ11において、インターネット70からHTMLを取得して、音声ブラウザクライアント13に転送し、当該音声ブラウザクライアント13において当該HTMLの解析を行う。また、Webブラウザ11は、HTMLに埋め込まれた音声データをインターネット70に要求し(ステップ203)、インターネット70から取得した当該要求に対応する応答を取得して、再生用ソフトウェア14を起動して転送する(ステップ204)。また、音声ブラウザクライアント13は、解析された結果に基づいて音声認識エンジン53に対して音声認識要求を発行し(ステップ205)、音声認識結果を取得すると、当該結果を音声ブラウザクライアント13に転送する(ステップ206)。これにより音声ブラウザクライアント13は再生用ソフトウェア14を制御して、音声再生する。

【0036】なお、本発明は、上記の実施例に限定されることなく、特許請求の範囲内で種々変更・応用が可能である。

#### 【0037】

【発明の効果】上述のように、本発明によれば、音声入力によってインターネットのWWW情報にアクセスし、音声で出力することが可能となる。また、アクセスした情報が英文の場合、翻訳機能によって日本語で出力可能であるので、英語の知識が少ないユーザでも、情報の内容を把握することができる。

【0038】さらに、テキスト文のみでなく、音楽情報、朗読情報へのアクセス可能であるため、娯楽用としても使用することができる。さらに、音楽情報や朗読情報の再生の際には、通常のラジカセ(出力・再生・録音機能を有するラジオ・カセット)と同様に、一時停止、再生、停止などの再生制御が可能であるため、ユーザの抵抗感が少ない。

【0039】さらに、テキスト音読中などに割り込んで時報を聞くことが可能であるため、視覚障害者向、または、時計を持ち合わせていない場合でも音声で時報を知ることができる。以上の機能により視覚障害者のインターネット利用への支援が可能となる。

#### 【図面の簡単な説明】

【図1】本発明の原理構成図である。

【図2】本発明が適用されるシステム構成図である。

【図3】本発明の一実施例のシステム構成図である。

【図4】本発明の一実施例の動作のシーケンスチャートである。

【図5】本発明の一実施例のWebブラウザの通常のユーザインタフェースを示す図である。

\*【図6】本発明の一実施例の音声ブラウザにより音読される例である。

【図7】本発明の一実施例のWebブラウザのマルチフレーム構成によって表示されているホームページの型の例である。

【図8】本発明の一実施例の音声ブラウザにより音読される例である。

【図9】本発明の一実施例のWebブラウザにイメージが掲載された例である。

10 【図10】本発明の一実施例の音声ブラウザにより音読される例である。

【図11】本発明の一実施例の音楽/朗読コンテンツ再生を行う場合の構成図である。

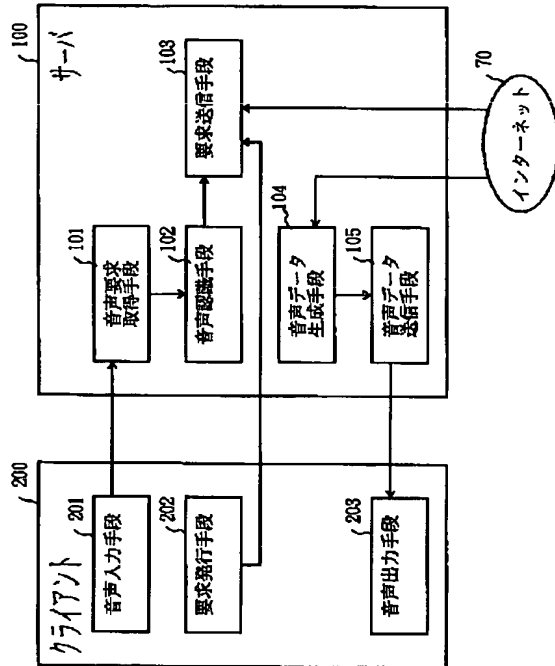
【図12】本発明の一実施例の音声データの再生制御のシーケンスチャートである。

#### 【符号の説明】

- 10 クライアント端末
- 11 Webブラウザ
- 12 音声入力用ボタン監視用プログラム
- 20 13 音声ブラウザクライアント
- 14 再生用ソフトウェア
- 20, 30, 40, 50, 60 ワークステーション
- 21 プロキシサーバ
- 22 英日翻訳エンジン
- 31 日英翻訳エンジン
- 41 プロキシサーバ
- 51 音声認識I/Fプログラム
- 52 音声認識エンジン
- 53 形態素解析エンジン
- 30 61 音声合成エンジン
- 70 インターネット
- 100 サーバ
- 101 音声要求手段
- 102 音声認識手段
- 103 要求送信手段
- 104 音声データ生成手段
- 105 音声データ送信手段
- 110 タイトル
- 120 本文
- 40 130 リンク項目
- 200 クライアント
- 201 音声入力手段
- 202 要求発行手段
- 203 音声出力手段
- 210 第一フレーム
- 220 第二フレーム
- 310 イメージ図

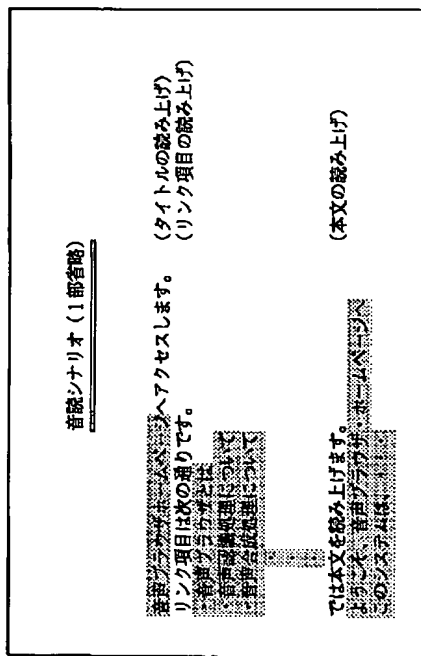
【図 1】

本発明の原理構成図



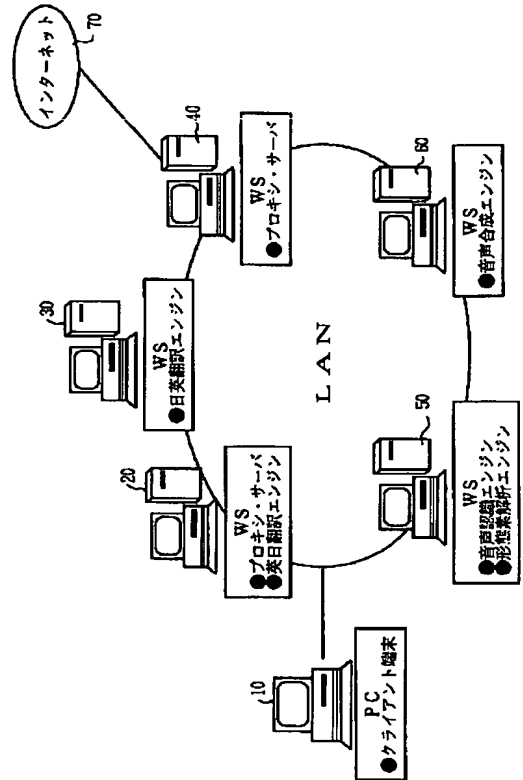
【図 6】

本発明の一実施例の音声ブラウザにより音読される例



【図 2】

本発明が適用されるシステム構成図



【図 9】

本発明の一実施例のWebブラウザにイメージが記載された例

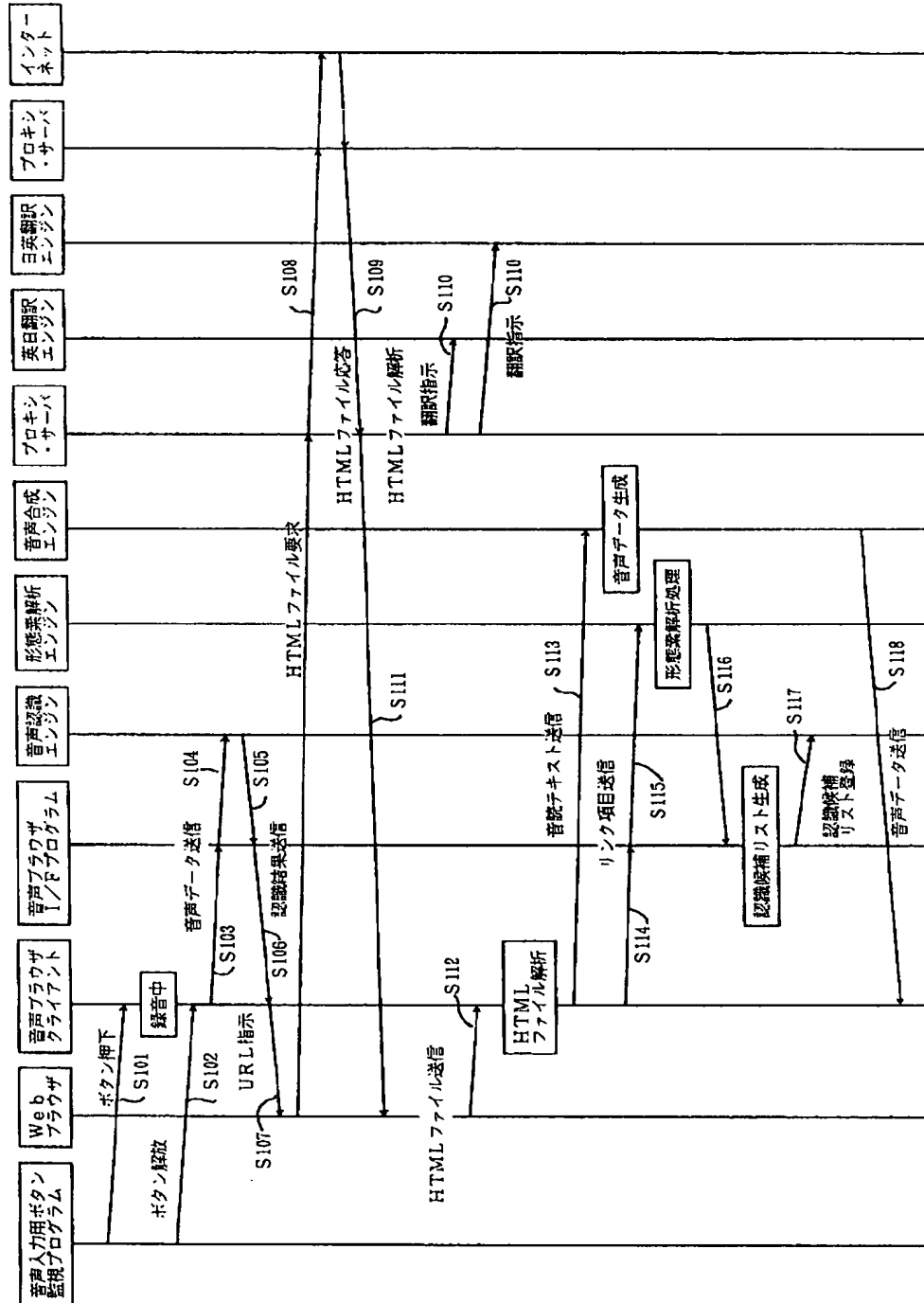






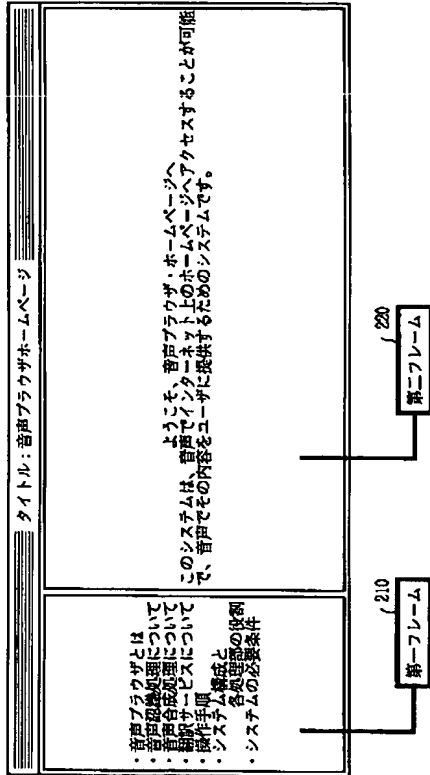
【図4】

## 本発明の一実施例の動作のシーケンスチャート



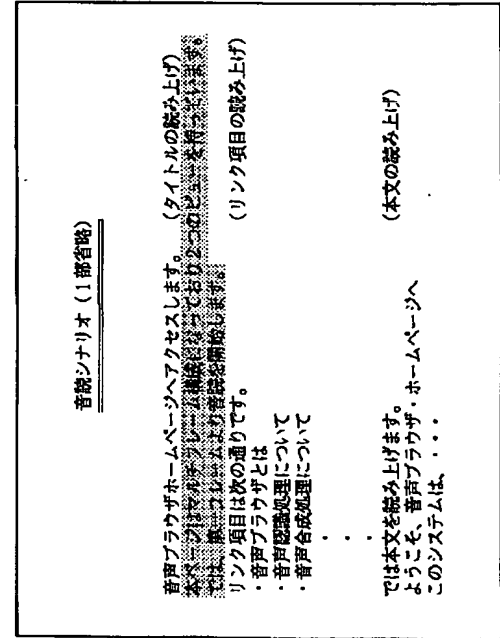
【図 7】

本発明の一実施例のWebブラウザマルチフレーム構成によって表示されているホームページの型の例



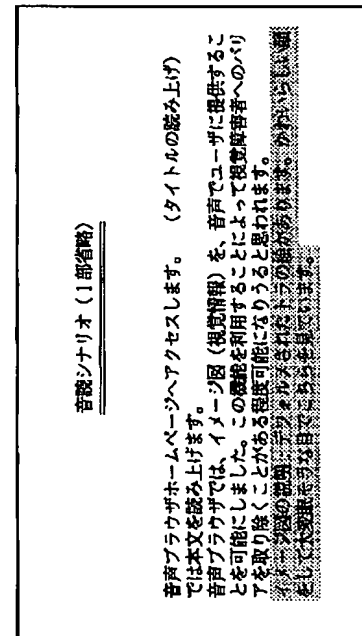
【図 8】

本発明の一実施例の音声ブラウザにより音読される例



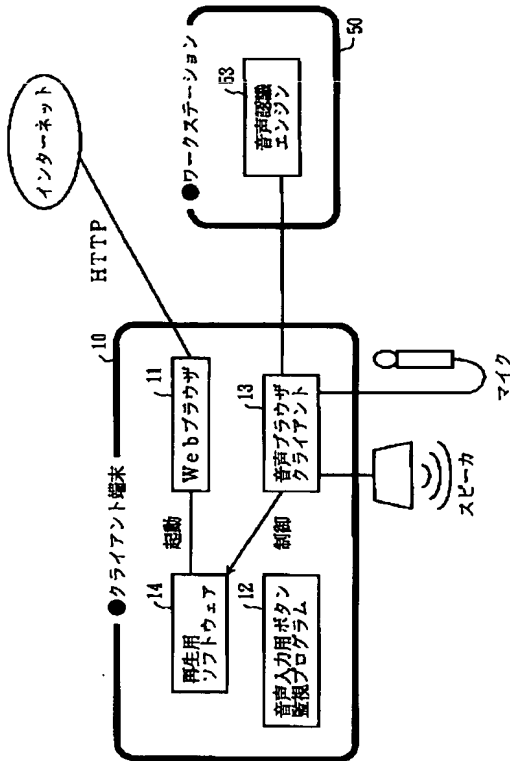
【図 10】

本発明の一実施例の音声ブラウザにより音読される例



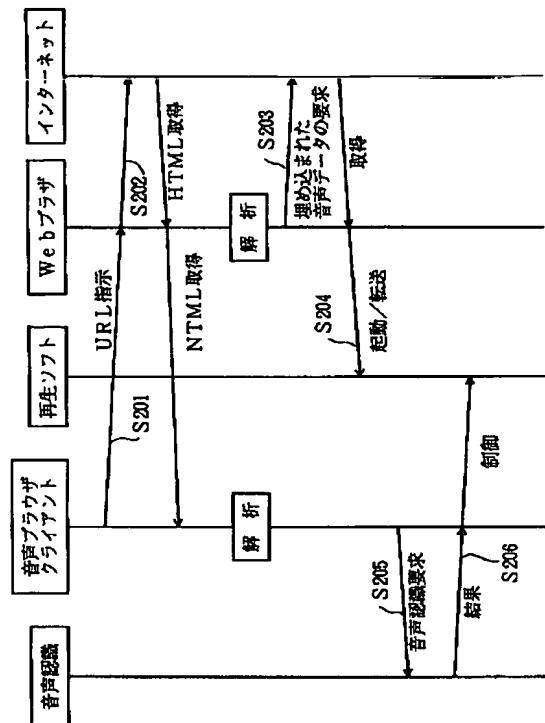
【図 1 1】

本発明の一実施例の音楽／朗読  
コンテンツ再生を行う場合の構成図



【図 1 2】

本発明の一実施例の音声データの  
再生制御のシーケンスチャート



フロントページの続き

(72)発明者 金丸 哲哉

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72)発明者 木目田 常治

東京都新宿区西新宿三丁目19番2号 日本  
電信電話株式会社内

(72)発明者 氏家 正美

東京都渋谷区桜丘町20番1号 エヌティ  
ィエレクトロニクス株式会社内